SERIE DE REVISION STATISTIQUE

EXERCICE N°1 : 20' 4 points

Dans le tableau statistique suivant, X désigne la température extérieure en 24 heures et Y désigne la consommation de pétrole de chauffage pour les même 24 heures et pour une famille donnée.

X en degrés	-2	0	4	8	10
Y en litres	40	30	20	15	10

1°) Construire dans un plan rapporté à un repère orthogonal le nuage de points représentant la série statistique à double donnée.
2°) Calculer les coefficients de corrélation linéaire entre X et Y et vérifier qu'il y a une forte corrélation linéaire entre ces deux variables.
3°) Déterminer une équation cartésienne de la droite de régression de Y en X.
4°) Quelle prévision (en litres) sur sa consommation de pétrole peut faire la famille considérée si une vague de froid persiste pendant 48 heurs avec une température moyenne extérieure de –4° ?

EXERCICE N°2 : 20' 4 points

Une urne contient 7 boules, constituées de différentes matières (fer, cuivre, aluminium),
Sont reparties suivant leur volume x exprimé en cm^3 et leur masse y exprimer en kg.

Volume x_i (cm^3)	10	12	14	16	19	20	23
Masse y_i (kg)	4	5	5	5	6	8	8

1) a- calculer les moyennes arithmétiques $\bar{x}$ et $\bar{y}$ ainsi les variances v(x) et v(y).
b- Calculer la covariance cov(x , y).
2) Calculer le coefficient de corrélation linéaire de cette série. Interpréter ce résultat.
3) a- Déterminer une équation cartésienne de la droite de régression de y en x.
b- Donner une estimation de la masse d'une boule de volume 30 cm^3 .
4) On tire au hasard et simultanément deux boules de l'urne.
a- Calculer la probabilité de l'événement S" obtenir deux boules de même masse "
b- On considère l'événement F" obtenir deux boules de volume inférieur ou égal à19 "
Calculer p(F) , $p(S \cap F)$ et p(S/F).
5) On répète l'épreuve précédente 4 fois de suite en remettant à chaque fois les boules dans l'urne.
Soit Z la variable aléatoire prenant pour valeur le nombre de fois ou S est réalisé .
a- Déterminer la loi de probabilité de Z.
b- Calculer l'espérance et la variance de Z.

EXERCICE N°3 : 20' 4 points

Une population de 40 propriétés agricoles a été étudier selon deux caractères :
X : le nombre de personnes actives et Y : la surface cultivable en hectares.
Les résultats de cette étude sont résumés dans le tableau statistique suivant :

Y \ X	1	2	3	4	5
[20 , 25 [	2	0	0	0	0
[25 , 30 [	5	1	0	0	0
[30 , 35 [	3	2	0	0	0
[35 , 40 [	1	7	1	0	0
[40 , 45 [	0	4	3	0	0
[45 , 50 [	0	1	4	1	0
[50 , 55 [	0	0	0	3	2

1°) Déterminer la distribution marginale de X. Quelle est le mode de cette série.
2°) Déterminer la distribution marginale de Y. Calculer la médiane de cette série.
3°) Représenter le nuage des points correspondant à cette série.
4°) Quelle conjecture vous propose la forme de ce nuage sur la corrélation des deux caractères ?
5°) Calculer les coordonnées du points moyen G.

EXERCICE N°4 : 30' 5 points

Le tableau suivant donne l'évolution du nombre d'adhérents d'un club de rugby de 2001 à 2006.

Année	2001	2002	2003	2004	2005	2006
Rang x_i	1	2	3	4	5	6
Nombre d'adhérents y_i	70	90	110	140	170	220

On cherche à étudier l'évolution du nombre y d'adhérents en fonction du rang x de l'année.

PARTIE A : un ajustement affine.

1) Représenter le nuage de points associé à la série (x_i,y_i) dans un repère orthogonal du Plan.
2cm pour une année sur $(o,\vec{i})$ et 1 cm pour 20 adhérents sur $(o,\vec{j})$.

2) a- Calculer v(x), v(y), cov(x,y).
b- Calculer le coefficient de corrélation linéaire de cette série. Interpréter ce résultat.

3) Déterminer une équation de la droite d'ajustement de y en x obtenue par la méthode des moindres carrés et la tracer sur le graphique précèdent (les coefficient arrondis à l'unité).

4) En supposant que cet ajustement reste valable pour les années suivantes, donner une estimation du nombre d'adhérents en 2009.

PARTIE B : un ajustement exponentiel.

On pose z = lny

1) Recopier le tableau suivant en arrondissant les valeurs de z_i au millième.

x_i	1	2	3	4	5	6
z_i	4.248					

2) Déterminer une droite d'ajustement de z en x par la méthode des moindres carrés (les coefficients seront arrondis au millième).

3) En déduire une approximation du nombre d'adhérents y en fonction du rang x de l'année.

4) En prenant l'approximation $y \approx 57.1e^{0.224x}$ et en supposant qu'elle reste valable pour les années suivantes, donner une estimation du nombre d'adhérents en 2009.

PARTIE C : comparaison des ajustements.

En 2009, il y a eu 430 adhérents. Lequel des deux ajustements semble le plus pertinent ? Justifier la réponse.